CHENG Sitong
https://cmots.github.io

EDUCATION

The Chinese University of Hong Kong 09/2021-11/2022
- **Master** of Science in Computer Science (*with distinction*)

Beijing University of Posts and Telecommunications 09/2017-07/2021
- Bachelor of Engineering, *Internet of Things Engineering*
- It's a joint program with Queen Mary University of London

Queen Mary University of London 09/2017-06/2021
- **Bachelor** of Science with Honors, *First Class*

WORK EXPERIENCE

Honorary Research Assistant in AIoT Lab, The Chinese University of Hong Kong 06/2022–12/2023
- Led a project in collaboration with MTR, using affordable scanning sonars and deep learning methods to detect drowning in swimming pools instead of camera. Responsible for hardware control, data preprocessing, model training, and real-time execution using Python and PyTorch.
- Responsible for applying data augmentation and SimCLR to improve model performance and robustness, addressing the challenge of low-resolution and limited availability of sonar data.
- Responsible for selectively skipping certain scan areas, then implementing Masked Autoencoders (MAE) to reconstruct images to compensate for the slow scanning speed (only 1 FPS).
- Achieved a fine-grained poses classification accuracy of 90.7% and 100% accuracy in binary safety/drowning classification during testing. Subsequently, a real-time prototype was successfully deployed at the Heng Fa Chuen residential swimming pool.

Full-time Intern in Intelligent Multimedia Group, Microsoft Research Asia, Beijing 02/2021-04/2021
- Working on keyword spotting, implemented some traditional transformer-related networks as baselines, and developed a pipeline to extract keywords from speech datasets.

PROJECT EXPERIENCE

LLM Prompt Recovery 03/2024-present
- Recovered the LLM (Large Language Model) prompt used for rewriting a given text. Employed LLM to predict prompts for rewriting, utilized known raw texts and rewritten texts generated by Gemma.
- Utilized QLoRa (4-bit quantization) to instruction-tune Mistral 7B on ~50M tokens of data. Reduced GPU memory usage through Flash Attention and DeepSpeed ZeRO3 (with CPU offload).
- Successfully improved the score from 0.54 to 0.65 now. Achieved a top 3% ranking in Kaggle out of approximately 2000 teams. Official Web: https://www.kaggle.com/competitions/llm-prompt-recovery

Machine Learning Technologies for Digital Biomarkers for Alzheimer's Disease 06/2022-12/2022
- Developed a sensor box containing multiple non-invasive sensors placed in living/bedroom rooms to capture various behavioral data like basic motion/social activities. The collected data is then analyzed using deep learning to identify potential signs and stages of Alzheimer's Disease.
- Responsible for assembling and installing hardware based on Raspberry Pi and Arduino to collect data of Alzheimer's disease.
- Responsible for implementing data preprocessing scripts to clean, standardize, and temporally align multi-source time series data from different sensors.

ASR-Free Pronunciation Assessment 12/2019-05/2020
- Investigated an ASR (automatic speech recognition)-free scoring approach that is derived from the marginal distribution of raw speech signals.
- Implemented some generative models to transfer raw speech signals to a featured distribution, including VAE and Normalizing flow.
- Responsible for programming, data collection and visualization.

PUBLICATIONS

- [Interspeech'20] **Cheng Sitong**., Liu, Z., Li, L., Tang, Z., Wang, D., Zheng, T.F.
  ASR-Free Pronunciation Assessment
  DOI: 10.21437/Interspeech.2020-2623.
- [MobiSys'23] Ouyang, Xiaomin and Xie, Zhiyuan and Fu, Heming and **Cheng, Sitong** and Pan, Li and Ling, Neiwen and Xing, Guoliang and Zhou, Jiayu and Huang, Jianwei
  Harmony: Heterogeneous Multi-Modal Federated Learning through Disentangled Model Training.
- [ICASSP'20] Fan, Yue and Kang, JW and Li, LT and Li, KC and Chen, HL and **Cheng, ST** and Zhang, PY and Zhou, ZY and Cai, YQ and Wang, Dong
  CN-Celeb: A Challenging Chinese Speaker Recognition Dataset